2Я

УДК 004.8/.93'1:519.254

А.Е. Янковская, С.В. Китлер
Томский государственный архитектурно-строительный университет, г. Томск, Россия
ayyankov@gmail.com, kitsv@sibmail.com

Принятие решений на основе параллельных алгоритмов тестового распознавания образов*

Статья посвящена принятию решений с использованием алгоритмов тестового распознавания образов в пространстве k-значных признаков. Описывается параллельный алгоритм построения безызбыточных безусловных диагностических тестов и алгоритм параллельного построения решающих правил, на основе которых принимается решение о принадлежности исследуемого (распознаваемого) объекта образу. Принятие решений осуществляется на базе логико-комбинаторных алгоритмов.

Введение

В области тестового распознавания образов используются различные методы и подходы решения задачи принятия решений [1-4]. Несмотря на то, что задача принятия решения рассматривалась в довольно большом количестве публикаций Ю.И. Журавлёва, Н.Г. Загоруйко, А.Д. Закревского, А.Е. Янковской, она до сих пор не потеряла своей актуальности. Решение данной задачи особенно важно для слабоструктурированных областей знаний, к которым относятся медицина, экология, геология, психология, социология.

При большом признаковом пространстве, как правило, число диагностических тестов, на основе которых принимается решение, весьма велико, что не позволяет принятие решения осуществлять в диалоговом режиме. В связи с этим стоит задача распараллеливания процесса как построения диагностических тестов, так и принятия решений на их основе. Нам неизвестны исследования в данном направлении. В связи с этим очевидна постановка и решение задачи разработки эффективных параллельных алгоритмов построения диагностических тестов и построения решающих правил, на основе которых принимается решение. При этом цель применения предложенной ранее процедуры голосования на множестве правил принятия решений [5], построенных на множестве диагностических тестов [4], [5] для повышения точности принимаемых решений, аналогична применению процедуры голосования, используемой в исследованиях Дж. Неймана и К. Шеннона в области теории автоматов, связанных с проблемами синтеза надежных машин из ненадежных элементов [6], [7].

Поскольку принятие решения о принадлежности исследуемого объекта к образу, как правило, не входящего в обучающую выборку, происходит в режиме реального времени, то для сокращения временных затрат предлагается алгоритм параллельного построения решающих правил и принятия решения на их основе.

В данном исследовании задача принятия решения осуществляется на основе k-значных диагностических тестов.

В настоящей статье описываются основные понятия и определения, приводится параллельный алгоритм построения диагностических тестов и алгоритм параллель-

*Работа выполнена при финансовой поддержке РФФИ (проект № 10-01-00462) и РГНФ (проект № 10-06-64604).

ного построения решающих правил, на основе которых осуществляется принятие решений с применением тестовых методов распознавания образов.

Целью данной работы является разработка эффективного алгоритма параллельного построения решающих правил в случае, когда количество используемых для принятия решений диагностических тестов значительно превышает количество процессоров, а также алгоритма параллельного принятия решений на основе тестовых методов распознавания образов.

Основные понятия и определения

Воспользуемся основными понятиями и определениями, введенными в публикациях [1], [4], [5] и применяемыми нами для дальнейшего изложения.

Для принятия решений на основе тестовых методов распознавания образов используется матричная модель представления данных и знаний [4], [5]. Поскольку данная модель применяется в основном в наших исследованиях, то приведём её описание.

Матричная модель включает целочисленную матрицу описаний (***Q***), задающую описание объектов в пространстве характеристических признаков $z_1, z_2, \ldots, z_m$ и целочисленную матрицу различений (***R***), задающую разбиение объектов на классы эквивалентности по каждому механизму классификации. Если значение характеристического признака несущественно для объекта, то данный факт отмечается прочерком («–») в соответствующем элементе матрицы ***Q***. Для каждого характеристического признака задаётся интервал изменения его значений. Множество всех неповторяющихся строк матрицы различений сопоставлено множеству выделенных образов, представленных одностолбцовой матрицей ***R'***, элементами которой являются номера образов.

На рис. 1 представим пример матричной модели данных и знаний.

Q=

z_1	z_2	z_3	z_4	z_5	z_6	z_7	
3	4	8	9	5	1	3	1
2	5	7	9	4	1	5	2
3	5	7	8	3	0	8	3
4	2	6	8	4	1	5	4
3	3	8	8	4	1	5	5
2	3	8	8	4	1	3	6
1	2	8	8	3	–	3	7
1	–	7	8	4	0	6	8
2	6	7	9	5	1	3	9
1	4	8	8	3	0	4	10

R=

k_1	k_2
1	1
1	1
1	2
1	2
1	1
2	1
1	2
2	2
2	1
2	2

R'=

1
1
2
2
1
3
2
4
3
4

Рисунок 1 – Матрицы ***Q***, ***R*** и ***R'***

Тестом называется совокупность признаков, различающих любые пары объектов, принадлежащих разным образам. Тест называется безызбыточным [4] (тупиковым [1]), если содержит безызбыточное количество признаков. Безызбыточный безусловный диагностический тест (ББДТ) характеризуется одновременным предъявлением всех входящих в него признаков исследуемого объекта при принятии решений. Безызбыточным столбцовым покрытием целочисленной матрицы называется подмножество столбцов, покомпонентная сумма элементов которых равна столбцу, состоящему из ненулевых значений, причем при исключении любого столбца из данного подмножества указанное свойство нарушается.

Предположим, что f-я строка целочисленной матрицы ***U*** поглощает t-ю строку ($U_f \succ U_t$)

$$\left(U_f \succ U_t\right) \leftrightarrow \forall i \in I\left(u_{fi} \geq u_{ti}\right), \qquad (1)$$

где $i \in \{1,2,\ldots,s\}$ – множество строк целочисленной матрицы, а u_{fi} (u_{ti}) – элемент матрицы ***U***, лежащей на пересечении f-й (t-й) строки и i-го столбца. Безызбыточной целочисленной матрицей назовем матрицу, в которой отсутствуют поглощающие строки.

Построение матрицы импликаций

Построение k-значных ББДТ сводится к построению матрицы импликаций [4], [5] – целочисленной матрицы $\boldsymbol{U}$, столбцы которой сопоставлены столбцам матрицы $\boldsymbol{Q}$, а строки – всевозможным парам объектов v, l из разных образов a, b соответственно (классов); $v\in\{1,2,...,\sigma(\boldsymbol{Q}_a)\}$, $l\in\{1,2,...,\sigma(\boldsymbol{Q}_b)\}$, где $\sigma(\boldsymbol{Q}_a)$ ($\sigma(\boldsymbol{Q}_b)$) – количество строк в подматрице $\boldsymbol{Q}_a$ ($\boldsymbol{Q}_b$). Подматрица $\boldsymbol{Q}_a$ ($\boldsymbol{Q}_b$) представляет собой целочисленную матрицу, строками которой являются строки матрицы $\boldsymbol{Q}$, принадлежащие образу a (b). Строка $\boldsymbol{U}_i$ матрицы $\boldsymbol{U}$ представляет собой значение целочисленной вектор-функции различения, j-я ($j\in\{1,2,...,m\}$) компонента u_{ij} которой вычисляется по формуле:

$$u_{i,j}=\left|q^{a}_{v,j}-q^{b}_{l,j}\right|, \qquad (2)$$

где $q^{a}_{v,j}$ ($q^{b}_{l,j}$) – значение признака z_j для объекта $\boldsymbol{v}$ (l), а u_{ij} вычисляется для каждой пары образов (классов при фиксированном механизме классификации).

Для матриц $\boldsymbol{Q}$ и $\boldsymbol{R'}$ (рис. 1) по формуле 1 построим матрицу импликаций.

$\boldsymbol{U}$=

	z_1	z_2	z_3	z_4	z_5	z_6	z_7
1	0	1	1	1	2	1	5
2	2	2	0	1	2	1	0
3	2	2	0	1	2	0	0
4	1	2	2	1	1	0	2
5	1	2	1	0	0	0	0
6	1	1	0	1	1	0	0
7	2	0	0	1	1	1	1
8	2	2	1	1	1	1	3
9	2	1	1	1	1	1	3
10	2	0	1	1	1	1	3
11	2	1	1	1	1	1	3
12	2	2	1	1	1	1	3
13	1	0	0	1	1	1	3
14	1	3	1	1	1	1	2
15	1	3	1	1	1	0	2
16	2	3	1	1	0	0	0
17	0	1	0	0	1	0	2
18	0	2	1	1	0	0	2
19	1	1	1	1	0	1	1
20	1	3	0	1	0	1	1
21	1	2	0	1	0	1	1
22	1	1	0	1	0	1	1
23	1	0	0	1	0	1	1
24	1	1	0	1	0	1	1
25	0	2	1	0	1	1	3
26	2	1	0	0	1	1	2
27	2	1	0	0	1	0	2
28	1	1	2	0	0	0	0
29	1	3	1	1	1	0	2
30	1	0	0	0	0	0	2
31	2	1	0	0	0	1	1
32	2	1	1	0	0	1	1
33	2	0	1	0	0	1	1
34	2	1	1	0	0	1	1
35	2	2	1	0	0	1	1
36	2	3	1	0	0	1	1
37	1	1	0	1	2	1	5
38	1	2	1	0	1	1	5
39	2	1	1	0	1	0	4
40	2	3	0	0	1	0	2
41	2	2	0	0	1	0	2
42	2	1	0	0	1	0	2
43	2	0	0	0	1	0	2
44	2	1	0	0	1	0	2
45	1	4	1	1	2	1	0
46	1	1	0	0	1	1	0
47	0	2	0	0	1	0	1
48	0	0	1	0	1	0	3
49	0	1	1	0	1	0	3
50	0	2	1	0	1	0	3
51	0	3	1	0	1	0	3
52	0	4	1	0	1	0	3
53	1	4	1	1	2	0	0
54	1	1	0	0	1	0	0
55	0	2	0	0	1	1	1
56	0	0	1	0	1	1	3
57	0	1	1	0	1	1	3
58	0	2	1	0	1	1	3
59	0	3	1	0	1	1	3
60	0	4	1	0	1	1	3
61	2	4	1	1	1	0	2
62	2	1	2	0	1	0	2
63	3	2	2	0	0	1	1
64	3	0	1	0	0	1	1
65	3	1	1	0	0	1	1
66	3	2	1	0	0	1	1
67	3	3	1	0	0	1	1
68	3	4	1	0	0	1	1
69	1	2	1	1	1	1	1
70	1	4	0	1	1	1	3
71	1	3	0	1	1	1	3
72	1	2	0	1	1	1	3
73	1	1	0	1	1	1	3
74	1	0	0	1	1	1	3
75	1	1	0	0	0	1	1
76	1	1	1	0	0	1	3
77	1	0	1	0	0	1	3
78	1	1	1	0	0	1	3
79	1	2	1	0	0	1	3
80	1	3	1	0	0	1	3

Рисунок 2 – Матрица $\boldsymbol{U}$

После удаления в матрице $\boldsymbol{U}$ (рис. 2) поглощающих строк (1) получим безызбыточную матрицу импликаций ($\boldsymbol{U'}$), представленную на рис. 3.

$$
\boldsymbol{U'} = \begin{array}{r} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \end{array}
\begin{array}{c} \begin{array}{ccccccc} z_1 & z_2 & z_3 & z_4 & z_5 & z_6 & z_7 \end{array} \\
\left[\begin{array}{ccccccc}
1 & 2 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 2 \\
0 & 2 & 1 & 1 & 0 & 0 & 2 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 \\
1 & 1 & 2 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 2 \\
2 & 0 & 1 & 0 & 0 & 1 & 1 \\
0 & 2 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 0 & 3 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 1
\end{array}\right] \end{array}
$$

Рисунок 3 – Матрица $\boldsymbol{U'}$

Закономерности в данных и знаниях

Под закономерностями [4] в знаниях будем понимать подмножества признаков с определенными легко интерпретируемыми свойствами, влияющими на различимость объектов из разных образов, устойчиво наблюдаемыми для объектов из обучающей выборки и проявляющимися на других объектах той же природы, а также весовые коэффициенты признаков.

К упомянутым подмножествам будем относить следующие подмножества признаков: константные, устойчивые (константные внутри образа), неинформативные (не различающие ни одной пары объектов из разных образов), альтернативные (одинаково различающие одни и те же объекты), квазиальтернативные (в разной степени различающие одни и те же объекты), зависимые (в смысле включения подмножеств различимых пар объектов), квазизависимые (в различной степени различающие различаемые объекты, причём квазизависимый признак различает меньшее число объектов, чем квазизависящий признак), несущественные (не входящие ни в один безызбыточный тест, используемый при распознавании для принятия решений), обязательные (входящие во все безызбыточные тесты), псевдообязательные (входящие в множество используемых при распознавании безызбыточных тестов и не являющиеся обязательными) признаки, а также все минимальные и все (либо часть – при большом признаковом пространстве) безызбыточные различающие подмножества признаков, являющиеся, по сути, соответственно минимальными и безызбыточными тестами, а также весовые коэффициенты признаков, основанные на их разделяющей способности.

Признаки, являющиеся квазиальтернативными (за исключением одного) и квазизависимыми, будем исключать при построении столбцовых покрытий матрицы импликаций, а в дальнейшем квазиальтернативные признаки будем использовать при построении решающих правил.

Для представления формулы вычисления значения весового коэффициента j-го признака (w_j) и формулы вычисления значения веса t-го теста (W_t) введём ряд обозначений:

w_j – весовой коэффициент целочисленного признака z_j ($j \in \{1,2,...,m\}$);

q_{aj}, q_{bj} – значение признака z_j для пары объектов из разных образов a, b ($a \neq b$);

δ_{ab}^{j} – величина j-й компоненты вектор-функции различения между объектами из разных образов a, b ($a \neq b$), вычисляемая по формуле: $\delta_{ab}^{j} = \left| q_{aj} - q_{bj} \right|$;

N_a – число строк в описании a-го образа;
K – число выделенных образов;
S_j – интервал изменения j-го признака в l-й строке матрицы $\boldsymbol{Q}$;
p_l – коэффициент повторения l-й строки;

$$d_{lj}^{-} = \begin{cases} 1, & если \quad q_{li} = "-" \\ 0, & если \quad q_{li} \neq "-" \end{cases}$$

σ_a – количество объектов в образе a (a = 1, ..., K), вычисляемое по формуле:

$$\sigma_a^j = \sum_{l=1}^{N_a} p_l \prod_{i}^{m} (S_i)^{d_{li}^{-}};$$

r, s ($r \neq s$) и a, b ($a \neq b$) – номера образов;
$N = \{N_1, N_2,\ldots, N_\tau\}$ – множество тестов;
L_t – множество признаков, входящих в тест N_t ($t \in \{1,\ldots,\tau\}$).

Весовой коэффициент w_j целочисленного признака z_j будем определять по его разделяющей способности аналогично вычислению w_j для двоичного признака z_j [4].

$$w_j = \frac{\sum_{r=1}^{K-1}\sum_{s=r+1}^{K}\sum_{a=1}^{N_r}\sum_{b=1}^{N_s}\delta_{ab}^{j}}{S_j\sum_{a=1}^{K-1}\sum_{b=a+1}^{K}\sigma_a\sigma_b}, \qquad (3)$$

Множество обязательных признаков будем называть ядром всех ДТ, поскольку исключение любого признака из ядра нарушает свойство каждого из тестов быть тестом.

Каждому тесту $N_t \in N$ соответствует вес теста W_t, вычисляемый по формуле:

$$W_t = \sum_{j \in L_t} w_j. \qquad (4)$$

Выявим закономерности по матрице $\boldsymbol{U'}$ (рис. 2).

В число закономерностей входят следующие подмножества признаков:
z_4 – зависимый признак от z_7;
z_6 – квазизависимый признак от z_7.

Для матрицы $\boldsymbol{U'}$ отсутствует ядро, а также другие подмножества признаков, входящие в закономерности.

Распознавание исследуемых объектов и принятие решений

Задача распознавания состоит в определении по матрицам $\boldsymbol{Q}$ и $\boldsymbol{R}$ образа, которому принадлежит заданный совокупностью признаков исследуемый объект, как правило, не входящий в обучающую выборку [8]. Для решения данной задачи необходимо построить все безызбыточные столбцовые покрытия матрицы $\boldsymbol{U'}$. Построение безызбыточных столбцовых покрытий целочисленной матрицы $\boldsymbol{U'}$ осуществляется по алгоритму, похожему на алгоритм построения безызбыточных столбцовых покрытий булевой матрицы, описанной в публикации [9].

По всем найденным безызбыточным столбцовым покрытиям матрицы $\boldsymbol{U'}$ строятся все k-значные ББДТ [10], а также находятся различного рода закономерности, в том числе весовые коэффициенты. Основная идея построения k-значных ББДТ [10] заключается в многоуровневом разбиении матриц $\boldsymbol{Q}$ и $\boldsymbol{R'}$ на подматрицы $\boldsymbol{Q_{i,j}}$ и $\boldsymbol{R'_{i,j}}$ и

параллельном построении подматриц $\boldsymbol{U'_{i,g}}$ по подматрицам $\boldsymbol{Q_{i,j}}$ и $\boldsymbol{R'_{i,j}}$ и объединением подматриц $\boldsymbol{U'_{i,g}}$ в матрице $\boldsymbol{U'}$ с одновременным удалением поглощающих строк. Процедура разбиения продолжается до тех пор, пока все строки подматрицы $\boldsymbol{Q_{i,j}}$ не будут принадлежать только одному образу.

Для описания параллельного алгоритма построения k-значных ББДТ введем следующие обозначения:

i – количество уровней разбиения матриц $\boldsymbol{Q}$ и $\boldsymbol{R'}$ ($i\in\{1, 2, ..., v\}$);

j – количество подматриц $\boldsymbol{Q_{i,j}}$ на i-м уровне разбиения ($j\in\{1, 2, ..., r_i\}$), j является функцией от i, т.е. $j(i)$;

g – количество подматриц $\boldsymbol{U'_{i,g}}$ на i-м уровне разбиения $g\in\{1, 2, ..., r_i/2\}$;

$]\log_2 f[$ – наименьшее сверху целое к $\log_2 f$;

f – количество образов в матрице $\boldsymbol{R'}$;

v – количество уровней разбиения, вычисляемое по формуле:

$v =]\log_2 f[$;

η – коэффициент, принимающий значение $\eta = 0$ при $i \neq v$ и $\eta = 1$ при $i = v$.

r_i – общее количество подматриц $\boldsymbol{Q_{i,j}}$ на i-м уровне разбиения, вычисляемое по формуле:

$$r_i = 2^i - 2\eta\left|f - 2^i\right|.$$

Пусть заданы матрицы $\boldsymbol{Q}$ и $\boldsymbol{R'}$ (рис. 1) и диапазон изменения значений следующих признаков: значение 1-го признака изменяется от 1 до 4, 2-го – от 2 до 6, 3-го – от 7 до 8, 4-го – от 8 до 9, 5-го – от 3 до 5, 6-го – от 0 до 1, 7-го – от 3 до 8.

Необходимо построить ББДТ.

Опишем краткий алгоритм параллельного построения k-значных ББДТ.

1. Упорядочивание строк матриц $\boldsymbol{Q}$ и $\boldsymbol{R'}$ по принадлежности к образам.

2. Разбиение матриц $\boldsymbol{Q}$ и $\boldsymbol{R'}$ на две пары подматриц $\boldsymbol{Q_{1,1}}$, $\boldsymbol{R'_{1,1}}$ и $\boldsymbol{Q_{1,2}}$, $\boldsymbol{R'_{1,2}}$ и дальнейшее разбиение подматриц $\boldsymbol{Q_{i,j}}$ и $\boldsymbol{R'_{i,j}}$ на подматрицы $\boldsymbol{Q_{i+1,j}}$, $\boldsymbol{R'_{i+1,j}}$ и $\boldsymbol{Q_{i+1,j+1}}$, $\boldsymbol{R'_{i+1,j+1}}$. Разбиение прекращается, когда в каждой паре подматриц $\boldsymbol{Q_{i,j}}$ и $\boldsymbol{R'_{i,j}}$ и $\boldsymbol{Q_{i,j+1}}$ и $\boldsymbol{R'_{i,j+1}}$ останутся строки, сопоставленные объектам, принадлежащим только одному образу.

3. На каждом i-м уровне разбиения по парам подматриц $\boldsymbol{Q_{i,j}}$ и $\boldsymbol{R'_{i,j}}$ и $\boldsymbol{Q_{i,j+1}}$ и $\boldsymbol{R'_{i,j+1}}$ осуществляется построение подматриц $\boldsymbol{U'_{i,g}}$ с применением формулы 2 с одновременным удалением поглощающих строк. Построение подматриц $\boldsymbol{U'_{i,g}}$ осуществляется параллельно на каждом процессоре.

4. Объединение всех подматриц $\boldsymbol{U'_{i,g}}$ в матрицу $\boldsymbol{U'}$ с удалением поглощающих строк, выявлением закономерностей и вычислением весовых коэффициентов признаков по формуле 3.

5. Построение всех безызбыточных столбцовых покрытий матрицы $\boldsymbol{U'}$ либо части при превышении заданной величины λ числа безызбыточных столбцовых покрытий в процессе их построения.

6. Построение по найденным безызбыточным столбцовым покрытиям матрицы $\boldsymbol{U'}$ всех k-значных ББДТ. Переход к пункту 8.

7. Доопределение матрицы $\boldsymbol{Q}$ и переход к пункту 2.

8. Конец.

По матрице $\boldsymbol{U'}$ (рис. 3) построим ББДТ:

$T_1 = \{z_1, z_7\}$; $T_2 = \{z_2, z_7\}$; $T_3 = \{z_1, z_2, z_3\}$; $T_4 = \{z_1, z_2, z_5\}$; $T_5 = \{z_3, z_5, z_7\}$.

Поскольку принятие решения осуществляется в реальном времени, то при большом признаковом пространстве для сокращения временных затрат предлагается алгоритм параллельного построения решающих правил, на основе которых прини-

мается решение о принадлежности исследуемого объекта образу (классу при фиксированном механизме классификации).

Сжатой матрицей описаний $\boldsymbol{Q}^i$ называется целочисленная матрица, построенная по матрице $\boldsymbol{Q}$, столбцы которой сопоставлены признакам, вошедшим в множество ББДТ, по которым принимается решение.

Далее приведём параллельный алгоритм построения решающих правил, который целесообразно применять при условии, что количество процессоров существенно меньше, чем количество построенных ББДТ.

1. Проверка матрицы $\boldsymbol{Q}$ на репрезентативность. В случае, если матрица $\boldsymbol{Q}$ не репрезентативна, переход к шагу 3.
2. Построение областей сходства по каждому образу.
3. Упорядочивание ББДТ по их мощности. Распределение всех ББДТ с учётом их мощности между имеющимся количеством процессоров.
4. Построение сжатых описаний образов (матрица $\boldsymbol{Q}^i$) ($i\in\{1,2,...,\tau\}$, где τ – число ББДТ) [9] на каждом процессоре согласно распределения ББДТ с одновременным построением коэффициентов сходства объектов внутри образа [8]. Каждая матрица $\boldsymbol{Q}^i$ строится из строк матрицы $\boldsymbol{Q}$ на основе признаков, вошедших в соответствующий распределению i-й ББДТ.
5. Удаление повторяющихся строк в матрицах $\boldsymbol{Q}^i$ (1) кроме одной из каждой группы повторяющихся строк.
6. Конец.

В результате построенное множество решающих правил реализует n независимых способов распознавания одного и того же объекта.

Распознавание исследуемого объекта по каждому тесту в рамках логико-комбинаторного подхода осуществляется с использованием порогового значения условной степени близости распознаваемого объекта к образам, вычисляемого на основе задаваемой пользователем (в процентах) допустимой погрешности распознавания.

Принятие решения о принадлежности исследуемого объекта образу осуществляется на основе процедуры голосования [5]. Если количество тестов соответствует количеству процессоров, то каждое принятие решений осуществляется на каждом процессоре. Иначе количество тестов равномерно распределяется между имеющимися процессорами и принятие решений осуществляется на каждом процессоре на основе распределённых ББДТ.

В целях изложения процедуры голосования введем следующие обозначения и понятия [5]. Число всех k-значных ББДТ, равное числу всех способов распознавания, обозначим через ρ. Через w_i^r обозначим вес r-го ББДТ, если распознаваемый объект принадлежит i-му образу при заданной погрешности распознавания ($r\in\{1,2, ..., \rho\}$, $i\in\{1, 2, ..., m\}$) и $w_i^r = 0$ в противном случае; через $\boldsymbol{w}$ – m-компонентный вектор, значение i-й компоненты которого равно сумме весов тестов при выводе о принадлежности объекта i-му образу по всем способам распознавания.

Алгоритм вывода о принадлежности распознаваемого объекта тому или иному образу осуществляется на основе процедуры голосования [5]:

1. $w := \{0, 0, ..., 0\}$;
2. $r :=1$;
3. i :=1;
4. $w := w + w_i^r$;
5. i := i +1; если i ≤ m, то переход к пункту 4;
6. r := r +1; если r ≤ ρ, то переход к пункту 3;

7. Определение номеров j, для которых $w_j \geq w_i$ $(i \neq j)$;

8. Принятие решения о принадлежности распознаваемого объекта образу j, если таковой один, и о равновероятной принадлежности образам с наибольшими равными величинами w_j;

9. Конец.

Итоговое решение о принадлежности объекта образу (классу при фиксированном механизме классификации) осуществляется на отдельном процессоре на основе процедуры голосования.

О реализации алгоритмов параллельного принятия решений в интеллектуальных системах

Принятие решений на основе параллельных алгоритмов построения k-значных ББДТ и построения решающих правил будет реализовано в интеллектуальном инструментальном средстве (ИИС) ИМСЛОГ [11], предназначенном для выявления различного рода закономерностей и принятия решений и являющемся интегрированной средой разработки прикладных интеллектуальных систем.

Архитектура ИИС ИМСЛОГ является открытой и представляет собой иерархическую систему программных модулей.

В настоящее время на базе ИИС ИМСЛОГ осуществляется разработка гибридной интеллектуальной системы диагностики и коррекции организационного стресса [8], [12], в которой будет реализовано принятие решений на основе параллельных алгоритмов тестового распознавания образов.

Выводы

Впервые поставлена задача принятия решений на основе параллельных алгоритмов тестового распознавании образов в пространстве k-значных признаков.

Описана используемая матричная модель представления данных и знаний и предложен эффективный параллельный алгоритм построения безызбыточных безусловных диагностических тестов.

Разработан эффективный параллельный алгоритм построения решающих правил и принятия решений на их основе.

Применение параллельных алгоритмов на основе тестовых методов распознавания образов позволит существенно сократить временные затраты при принятии решения об отнесении исследуемого (распознаваемого) объекта к образу.

Предложенные алгоритмы будут программно реализованы в ИИС ИМСЛОГ.

Дальнейшие исследования связаны с апробацией алгоритмов на реальных данных, с разработкой параллельного алгоритма обхода по ветвям дерева поиска, используемого при нахождении кратчайших, безызбыточных и оптимальных столбцовых покрытий целочисленной матрицы, а также с обеспечением отказоустойчивости безызбыточных безусловных диагностических тестов [13].

Литература

1. Журавлев Ю.И. Распознавание образов и анализ изображений / Ю.И. Журавлев, И.Б. Гуревич // Искусственный интеллект : в 3 кн. Книга 2. Модели и методы : справочник / под ред. Д.А. Поспелова. – М. : Радио и связь, 1990. – С. 149-191.
2. Загоруйко Н.Г. Прикладные методы анализа данных и знаний / Загоруйко Н.Г. – Новосибирск : ИМ СО РАН, 1999. – 270 с.

3. Закревский А.Д. Логика распознавания / Закревский А.Д. – Изд. 2-е, доп. – Москва : Едиториал УРСС, 2003. – 144 с.
4. Янковская А.Е. Логические тесты и средства когнитивной графики в интеллектуальной системе / А.Е. Янковская // Новые информационные технологии в исследовании дискретных структур : доклады 3-й Всероссийской конференции с международным участием. – Томск : Изд-во СО РАН. – 2000. – С. 163-168.
5. Янковская А.Е. Построение k-значных диагностических тестов в интеллектуальной системе с матричным представлением знаний / А.Е. Янковская // Сборник научных Трудов VI Национальной конференции по искусственному интеллекту с международным участием. – Пущино. – 1998. – Т. I. – С. 264-269.
6. J. von Neumann. Probabilistic logic and synthesis of reliable organisms from unreliable components / J. von Neumann // Automata Studies / edited by C. Shannon and J. McCarthy Princeton University Press, 1956. (Русский перевод: Автоматы. Москва : ИЛ, 1956. – С. 68-139).
7. Shannon C. Von Neumann's contributions to automata theory / C. Shannon // Bulletin of the American Mathematical Society. – 1958. – Vol. 64, № 2. – 123 p.
8. Янковская А.Е. Логические тесты и средства когнитивной графики в комплексе интеллектуальных систем : дис. ... доктора технических наук в форме научного доклада : 05.13.18 / А.Е. Янковская. – Томск, 2001. – 85 с.: ил. РГБ ОД, 71 02-5/211-1
9. Yankovskaya A.E. Construction and Evaluation of Compressed Descriptions of Patterns in an Intelligent Recognizing System / A.E. Yankovskaya and A.I. Gedike // Pattern Recognition and Image Analysis. – 1999. – Vol. 9, № 1. – P. 124-127.
10. Янковская А.Е. Основы построения гибридной интеллектуальной системы диагностики и коррекции организационного стресса с применением GRID технологии / А.Е. Янковская, С.В. Китлер // Международная научно-техническая конференция «Суперкомпьютерные технологии: разработка, программирование, применение» (СКТ-2010). – 2010 (в печати).
11. IMSLOG-2002 Software Tool for Supporting Information Technologies of Test Pattern Recognition / A.E. Yankovskaya, A.I. Gedike, R.V. Ametov, A.M. Bleikher // Pattern Recognition and Image Analysis. – 2003. – Vol. 13, № 4. – P. 650-657.
12. Янковская А.Е. Основы построения гибридной интеллектуальной системы диагностики и коррекции организационного стресса / А.Е. Янковская, Н.В. Казанцева, С.В. Китлер // Искусственный интеллект. Интеллектуальные системы (ИИ-2009). Материалы X Международной научно-технической конференции. – Таганрог : Изд-во ТТИ ЮФУ, 2009. – С. 130-133.
13. Янковская А.Е. Принятие решений, устойчивых к ошибкам измерения значений признаков в интеллектуальных системах / А.Е. Янковская // Искусственный интеллект. Интеллектуальные системы (ИИ-2009). Материалы X Международной научно-технической конференции. – Таганрог : Изд-во ТТИ ЮФУ, 2009. – С. 127-129.

А.Є. Янковська, С.В. Кітлер

Прийняття рішень на основі паралельних алгоритмів тестового розпізнавання образів

Стаття присвячена прийняттю рішень з використанням алгоритмів тестового розпізнавання образів у просторі k-значних ознак. Описується паралельний алгоритм побудови безнадлишкових безумовних діагностичних тестів і алгоритм паралельної побудови вирішальних правил, на основі яких приймається рішення про приналежність досліджуваного (того, що розпізнається) об'єкта образу. Прийняття рішень здійснюється на базі логіко-комбінаторних алгоритмів.

A.E. Yankovskaya, S.V. Kitler

Decision-Making on the Asis of Parallel Algorithms of Test Pattern Recognition

The paper is devoted to concerns decision-making with usage of algorithms of test pattern recognition in k-valued features space. The parallel algorithm of unconditional irredundant diagnostic tests construction and parallel algorithm of decision rules construction on the basis of which are made decision on membership of object under investigation to pattern are described. Decision-making is carried out on the base of logic-combinatorial algorithms.

Статья поступила в редакцию 02.07.2010.